

# Rationalizing Evaluativity

Dylan Bumford and Jessica Rett

UCLA

August 20, 2020

## evaluativity: an overview

a construction is **evaluative** iff it implies that some degree exceeds a contextual standard.

- canonical evaluativity: positive constructions like *Jane is tall*.
  - canonical non-evaluativity: explicit comparatives like *Jane is taller than Keisha*.
  - the Bierwisch (1989) Test: evaluative constructions entail the negation of their antonymic positive counterpart
- (1) a. Jane is tall.  $\rightarrow$  Jane is not short.  
b. Jane is taller than Keisha.  $\nrightarrow$  Jane is not short.

## evaluativity: the challenge

- the problem (Kamp 1975, Cresswell 1976): if the meaning of a comparative is compositionally derived in part from that of a positive construction, why isn't the comparative evaluative too?
- the canonical solution: a null operator POS which contributes evaluativity only in the absence of overt degree morphology

- (2) a.  $\llbracket \text{tall} \rrbracket^w = \lambda x. \mathbf{height}_w(x)$   
b.  $\llbracket \text{Jane is POS tall} \rrbracket^w = \mathbf{height}_w(j) \geq s$   
c.  $\llbracket \text{Jane is taller than Keisha} \rrbracket^w = \mathbf{height}_w(j) > \mathbf{height}_w(k)$

## Lassiter & Goodman (2013)

- instead, Lassiter & Goodman (2013) argue the evaluativity of positive constructions can be derived from pragmatic reasoning effects
    - gradable adjectives like *tall* denote relations between individuals, degrees
- (3) a.  $\llbracket \text{tall} \rrbracket^w = \lambda d \lambda x. \mathbf{height}_w(x) \geq d$   
b.  $\llbracket \text{Jane is } \emptyset_d \text{ tall} \rrbracket^w = \mathbf{height}_w(j) \geq d$
- in positive constructions, the degree argument is unsaturated
  - evaluativity emerges when listeners are forced to estimate a value for the degree argument  $d$  in addition to estimating a subject's height  $w$
  - they use a Rational Speech Act model (Frank & Goodman 2012)
    - not knowing which threshold  $d$  the speaker has in mind, the listener assumes the speaker has made a *rational choice* to utter the sentence
      - the speaker must think the sentence is reasonably informative (worth saying)
      - but at the same time, highly informative sentences are relatively likely
    - so the listener takes a weighted average of hypotheses about possible values the speaker may have had in mind, resulting in a degree argument value that is higher (but not by much) than the relevant standard

## evaluativity: a typology

- the L&G and POS approaches both assume that evaluativity surfaces in the absence of overt degree morphology
- but there is in fact no such correlation (Bierwisch 1989, Rett 2015)
- what's more, evaluativity can depend on antonymy

### (4) **non-evaluativity**

- a. Jane is 5 ft. tall. *measure phrase*
- b. Jane is taller/shorter than Keisha. *comparative*

### (5) **antonym-insensitive evaluativity**

- a. Jane is tall/short. *positive construction*
- b. Is Jane tall/short? *polar degree question*

### (6) **antonym-sensitive evaluativity**

- a. How tall/short is Jane? *degree question*
- b. Jane is as tall/short as Keisha. *equative*
- c. Jane is that tall/short too. *degree demonstrative*

## evaluativity: a typology

- the L&G and POS approaches both assume that evaluativity surfaces in the absence of overt degree morphology
- but there is in fact no such correlation (Bierwisch 1989, Rett 2015)
- what's more, evaluativity can depend on antonymy

### (4) **non-evaluativity**

- |                                        |                       |
|----------------------------------------|-----------------------|
| a. Jane is 5 ft. tall.                 | <i>measure phrase</i> |
| b. Jane is taller/shorter than Keisha. | <i>comparative</i>    |

### (5) **antonym-insensitive evaluativity**

- |                        |                              |
|------------------------|------------------------------|
| a. Jane is tall/short. | <i>positive construction</i> |
| b. Is Jane tall/short? | <i>polar degree question</i> |

### (6) **antonym-sensitive evaluativity**

- |                                     |                             |
|-------------------------------------|-----------------------------|
| a. How tall/short is Jane?          | <i>degree question</i>      |
| b. Jane is as tall/short as Keisha. | <i>equative</i>             |
| c. Jane is that tall/short too.     | <i>degree demonstrative</i> |

## the open argument problem

- in RSA, the listener's reasoning is driven by uncertainty
- but, in the traditional degree-semantic account (and L&G's), when a speaker asserts *Jane is as short as Keisha*, or *Jane is that short*, there isn't anything uncertain to reason about

(7) a.  $\llbracket \text{Jane is (exactly) that}_c \text{ short} \rrbracket^w =$

$$\mathbf{ht}_w(j) \underset{(\text{=})}{\leq} d_c$$

b.  $\llbracket \text{Jane is (exactly) as short as Keisha} \rrbracket^w =$

$$\{d : \mathbf{ht}_w(j) \leq d\} \underset{(\text{=})}{\supseteq} \{d : \mathbf{ht}_w(k) \leq d\}$$

- these adjectival constructions have degree arguments, like *Jane is tall* does, but those arguments are bound or valued overtly

## markedness in RSA

- the evaluativity contrast in (8) is due to the **relative markedness** of the negative adjective (Lehrer 1985, Rett 2015)

(8) a. Jane is (exactly) as tall as Keisha. *non-eval. equative*  
b. Jane is (exactly) as short as Keisha. *eval. equative*

- in fact, RSA models of markedness-driven Manner implicature have been proposed
- Bergen et al. 2016 (see also Potts et al. 2016) model the difference between synonymous short (i.e. unmarked) and long (i.e. marked) messages using a similar paradigm centered around lexical uncertainty



## Bergen et al. 2016

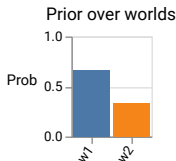
- two messages are denotationally equivalent, but one is more costly

$$\llbracket \text{marked} \rrbracket = \{w_1, w_2\} \quad \llbracket \text{unmarked} \rrbracket = \{w_1, w_2\}$$

- listener is uncertain of exactly what propositions the messages express, so they consider various *strengthenings* of the literal meaning

	$\llbracket \text{marked} \rrbracket$	$\llbracket \text{unmarked} \rrbracket$
$\mathcal{L}_0$	$\{w_1, w_2\}$	$\{w_1, w_2\}$
$\mathcal{L}_1$	$\{w_1\}$	$\{w_1, w_2\}$
$\mathcal{L}_2$	$\{w_1, w_2\}$	$\{w_1\}$
$\mathcal{L}_3$	$\{w_1\}$	$\{w_2\}$
$\vdots$	$\vdots$	$\vdots$

- some of these possible denotations are *a priori* more likely than others



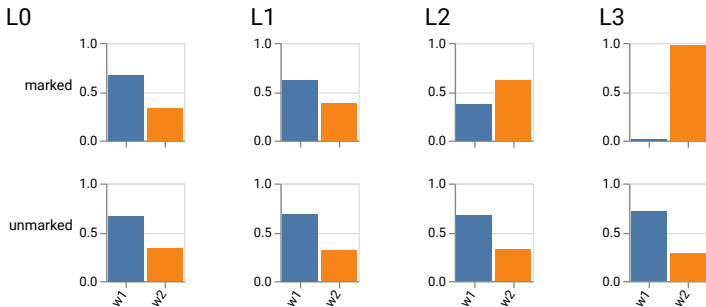
# Bergen et al. 2016

- listeners interpret utterances based on their prior beliefs and speakers' choices
- speakers choose utterances based on their cost and listeners' interpretations

$$\mathbb{L}_n(w | u, \mathcal{L}) \propto P(w) \cdot \mathbb{S}_n(u | w, \mathcal{L}) \quad [\text{greatly simplifying}]$$

$$\mathbb{S}_n(u | w, \mathcal{L}) \propto \mathbb{L}_{n-1}(w | u, \mathcal{L}) - C(u)$$

- under these conditions, marked messages are rationally interpreted as describing less likely scenarios, despite literal equivalence



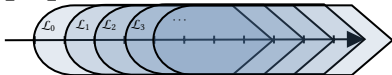
## LU doesn't help with the open argument problem

- does this help L&G with the anytonym-sensitive evaluativity contrast?

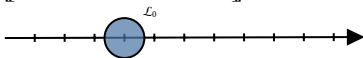
- (9) a. Jane is (exactly) as tall as Keisha. *non-*eval.* equative*  
b. Jane is (exactly) as short as Keisha. *eval. equative*

- these messages are plausibly denotationally equivalent, and (9b) more marked than (9a)
- but as things stand, they're both *maximally informative* with respect to the parameters under discussion; consider:

- $\llbracket \text{tall} \rrbracket =$



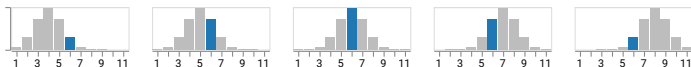
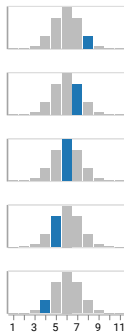
- $\llbracket \text{as tall/short as Keisha} \rrbracket =$



- so there are no strengthenings to consider; the use of the negative antonym in (9a) is simply inexplicable

# the new normal

- if the use of *short* (rather than *tall*) introduces evaluativity in certain constructions, there must be more to think about in these messages than the subject's height
- we propose that, in addition to reasoning about an adjective's degree argument, listeners reason about the *distribution* of that value within the relevant comparison class (Barker 2002)
  - an individual's height may be unknown within a known distribution
  - an individuals' height may be known within an unknown distribution

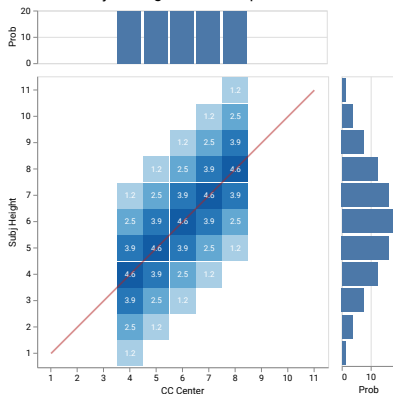


- in other words, there are many ways to be *tall*, and even many ways to be *5 ft tall*

# model assumptions, priors

- to model this, we assume
  - height is known to be normally distributed, though the center of the distribution is unknown
  - worlds are thus distinguished by both the height of the subject and the center of the comparison class
  - worlds where the subject's height is far from the mean are *a priori* unlikely

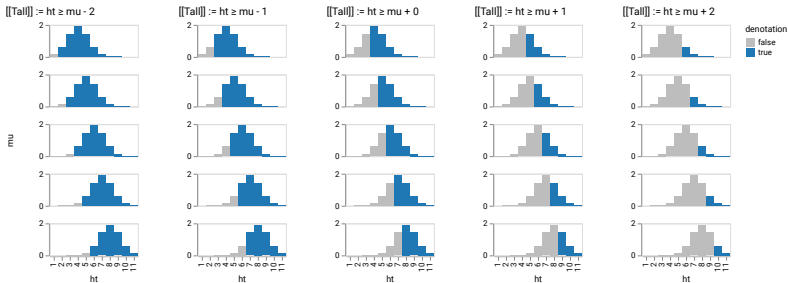
Prior over subject heights and comparison classes



# positive construction: semantic assumptions

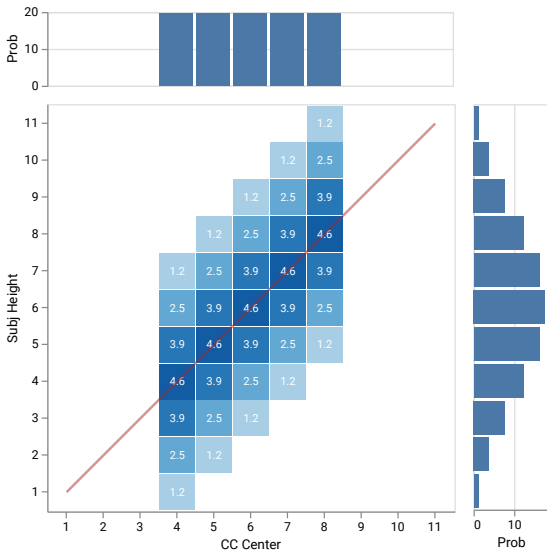
- positive constructions are tautologies  $\approx$  *Jane has a height* (Rett 2015)
- strengthened interpretations place subject's height in various upper (resp. lower) percentile of comparison class

	[[Jane is tall]]	[[Jane is short]]
⋮	⋮	⋮
$\mathcal{L}_{-1}$	$\lambda w. \mathbf{ht}_w(j) \geq \mu_w - 1$	$\lambda w. \mathbf{ht}_w(j) \leq \mu_w - 1$
$\mathcal{L}_0$	$\lambda w. \mathbf{ht}_w(j) \geq \mu_w + 0$	$\lambda w. \mathbf{ht}_w(j) \leq \mu_w + 0$
$\mathcal{L}_1$	$\lambda w. \mathbf{ht}_w(j) \geq \mu_w + 1$	$\lambda w. \mathbf{ht}_w(j) \leq \mu_w + 1$
⋮	⋮	⋮



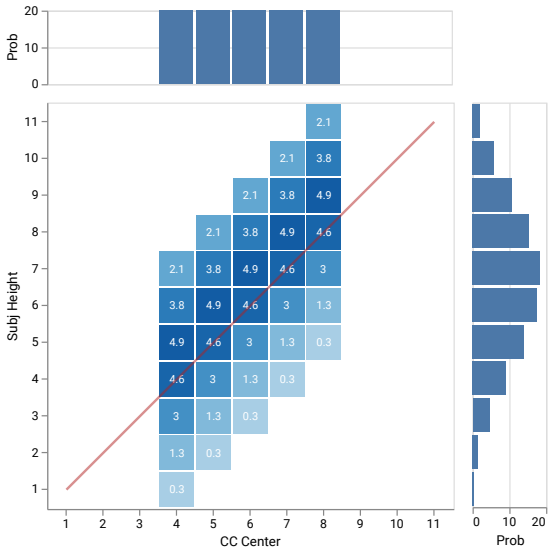
# positive simulation: *Jane is tall*

L0 literal listener



# positive simulation: *Jane is tall*

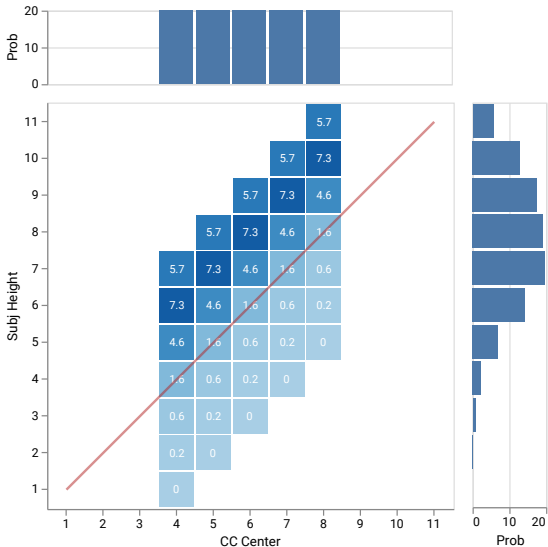
L0 + lex. uncertainty





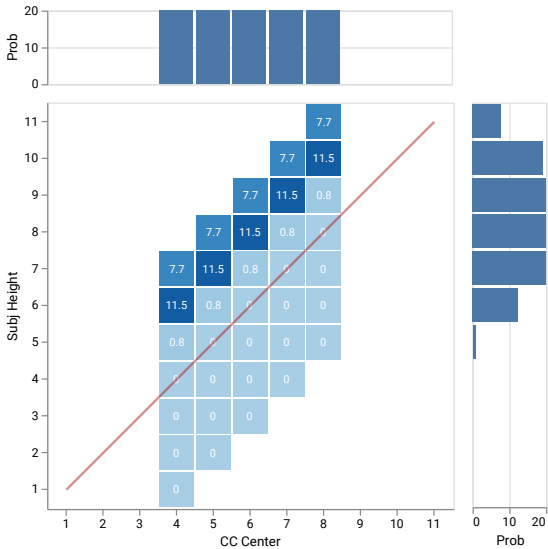
# positive simulation: *Jane is tall*

## L1 pragmatic listener



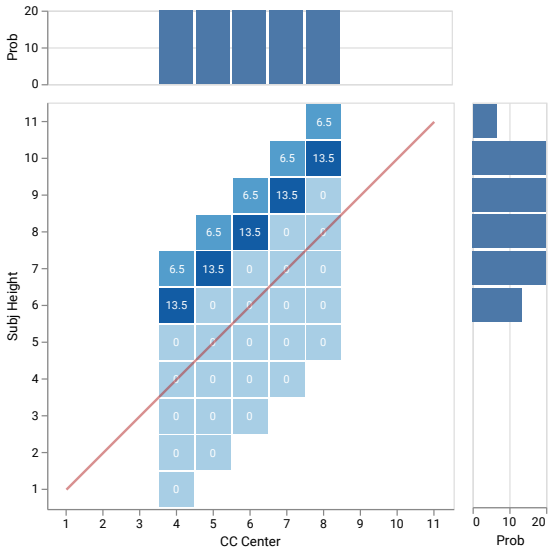
# positive simulation: *Jane is tall*

## L2 pragmatic listener



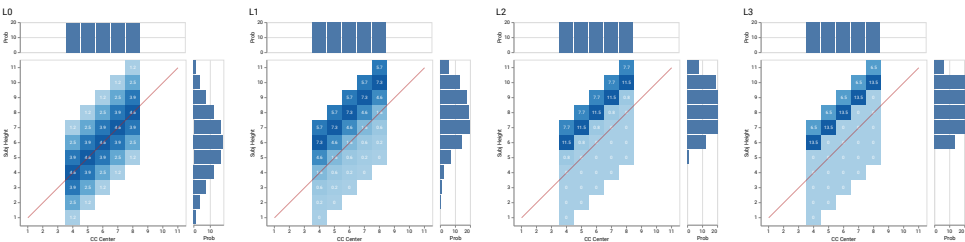
# positive simulation: *Jane is tall*

## L3 pragmatic listener



# positive construction simulation summary: *Jane is tall*

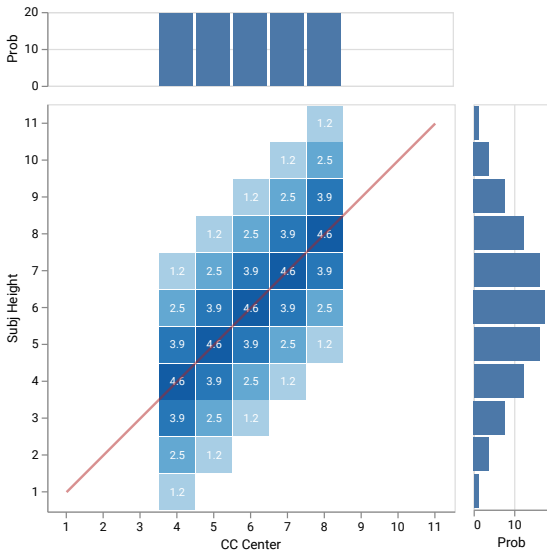
- As pragmatic iterations proceed, listeners become very confident that Jane's height exceeds the median, though they remain in the dark about what the median is



- And as seen in the height marginal, even though the standard remains unknown, the belief that Jane's height exceeds it leads to a (positively skewed) evaluative distribution over worlds
- evaluativity: **strong**

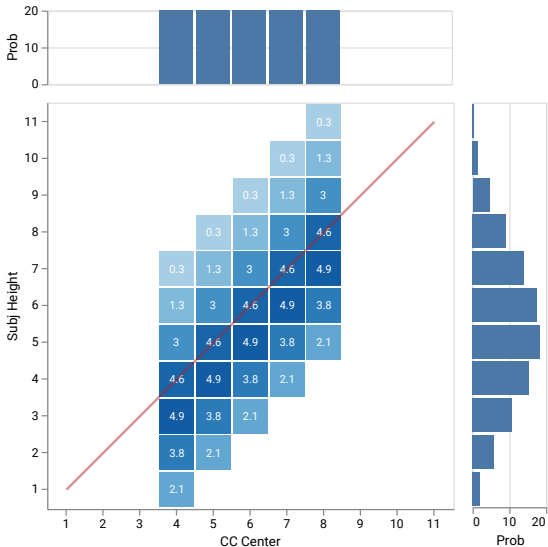
# positive construction simulation: *Jane is short*

L0 literal listener



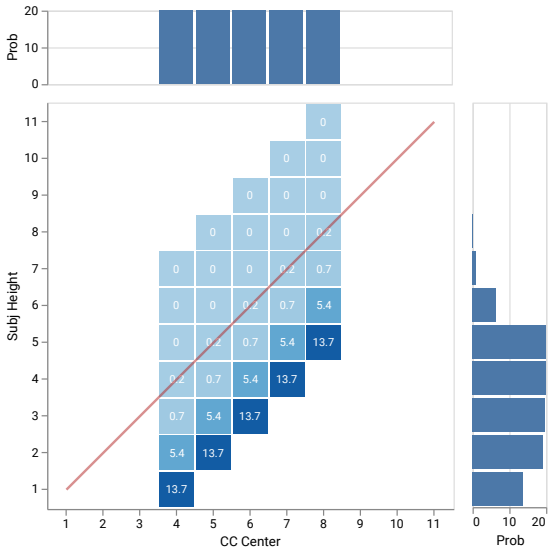
# positive construction simulation: *Jane is short*

L0 + lex. uncertainty



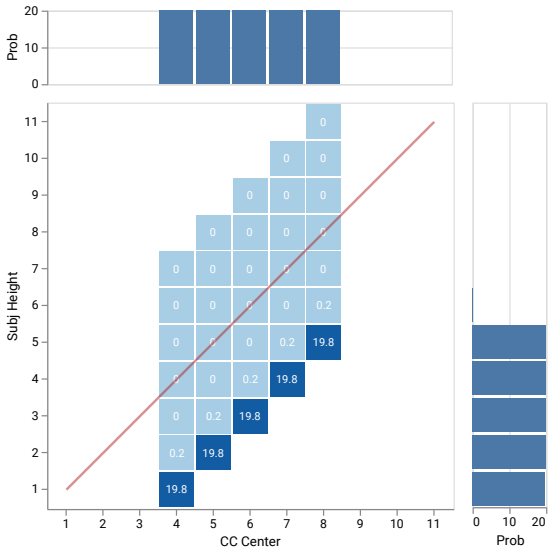
# positive construction simulation: *Jane is short*

## L1 pragmatic listener



# positive construction simulation: *Jane is short*

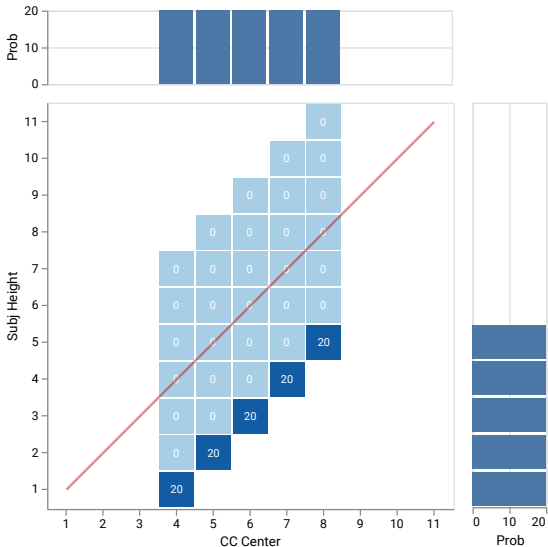
## L2 pragmatic listener





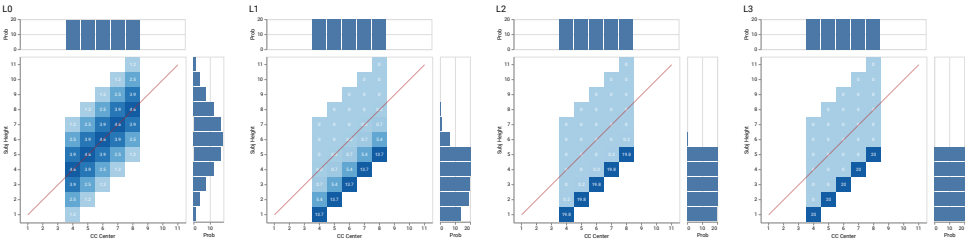
# positive construction simulation: *Jane is short*

## L3 pragmatic listener



# positive construction simulation summary: *Jane is short*

- As pragmatic iterations proceed, listeners become absolutely certain that Jane's height falls below the standard, though they remain in the dark about what the standard is



- Even though the center remains unknown, the belief that Jane's height falls below it leads to a (negatively skewed) evaluative distribution over worlds
- evaluativity: **very strong**

# degree demonstrative: semantic assumptions

- *that tall/short* both taken to denote that subject's height is exactly 6
- strengthened interpretations again place subject's height within various upper/lower percentiles of comparison class

[[Jane is that<sub>6</sub> tall]]

[[Jane is that<sub>6</sub> short]]

⋮

⋮

⋮

$$\mathcal{L}_{-1} \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \geq \mu_w - 1 \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \leq \mu_w - 1$$

$$\mathcal{L}_0 \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \geq \mu_w + 0 \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \leq \mu_w + 0$$

$$\mathcal{L}_1 \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \geq \mu_w + 1 \quad \lambda w. \mathbf{ht}_w(j) = 6 \wedge 6 \leq \mu_w + 1$$

⋮

⋮

⋮

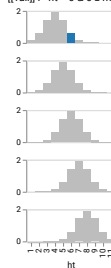
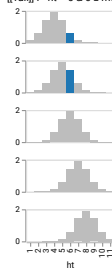
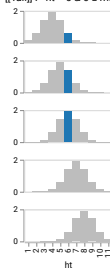
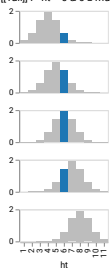
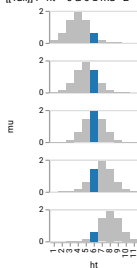
[[Tall]] := ht = 6 & 6 ≥ mu - 2

[[Tall]] := ht = 6 & 6 ≥ mu - 1

[[Tall]] := ht = 6 & 6 ≥ mu + 0

[[Tall]] := ht = 6 & 6 ≥ mu + 1

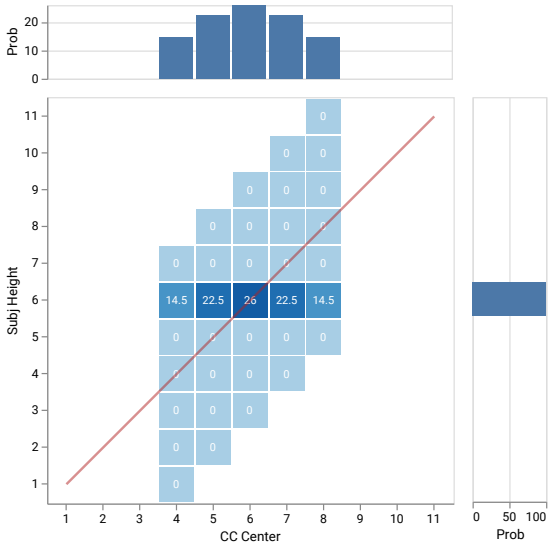
[[Tall]] := ht = 6 & 6 ≥ mu + 2



denotation  
■ false  
■ true

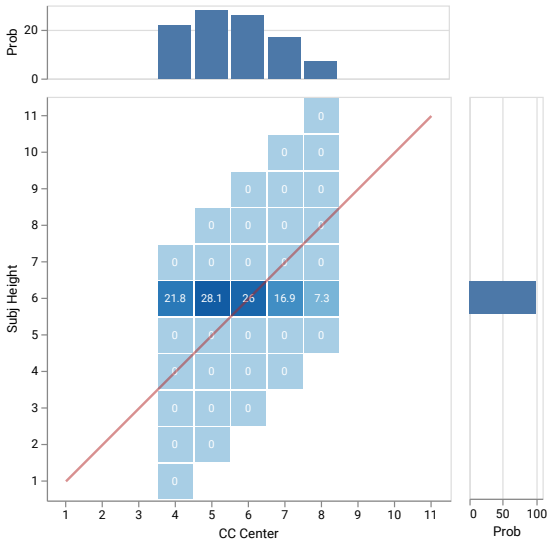
# degree demonstrative simulation: *Jane is that tall*

L0 literal listener



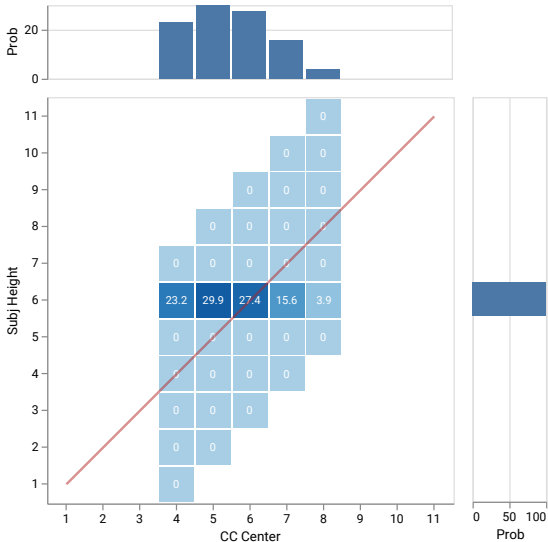
# degree demonstrative simulation: *Jane is that tall*

L0 + lex. uncertainty



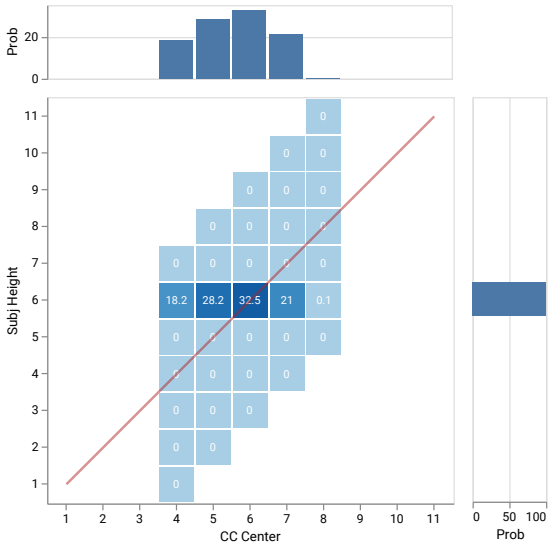
# degree demonstrative simulation: *Jane is that tall*

L1 pragmatic listener



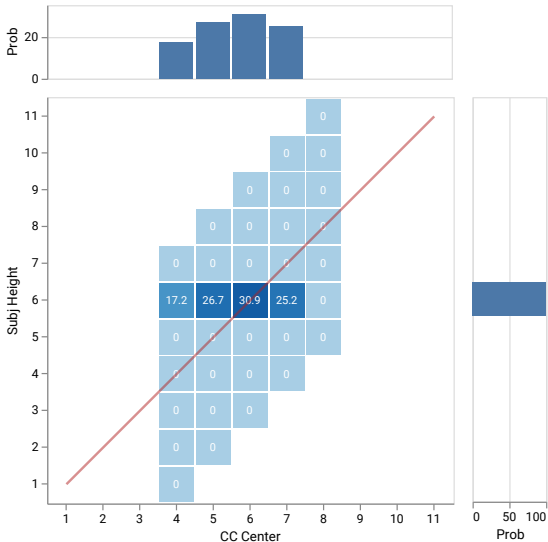
# degree demonstrative simulation: *Jane is that tall*

L2 pragmatic listener



# degree demonstrative simulation: *Jane is that tall*

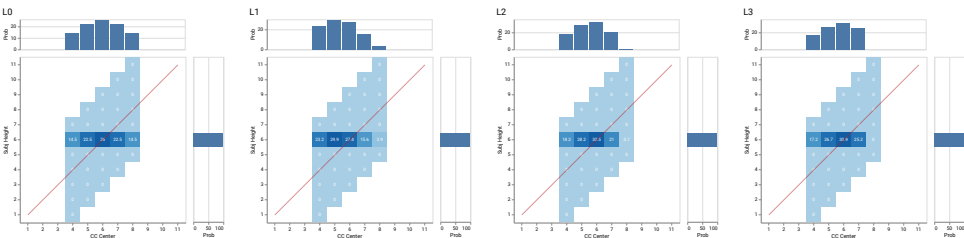
L3 pragmatic listener





## degree demonstrative summary: *Jane is that tall*

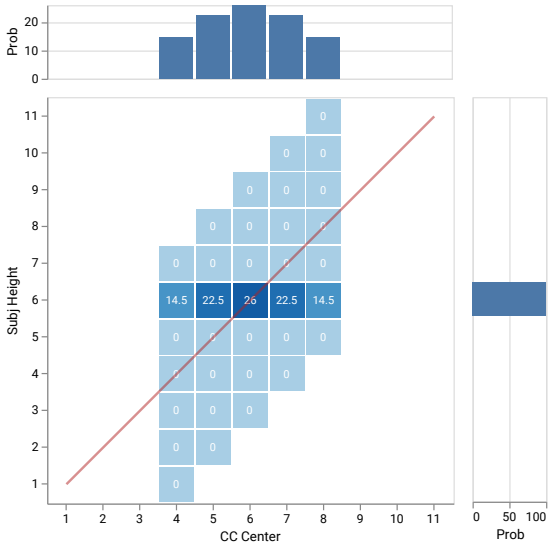
- Jane's height is of course fixed by the utterance, and since people are usually normally-sized, a literal interpreter will assume 6 is the most likely center



- As iterations proceed, the listener's belief does not move far from where it is after interpreting the sentence literally
- evaluativity: **very weak**

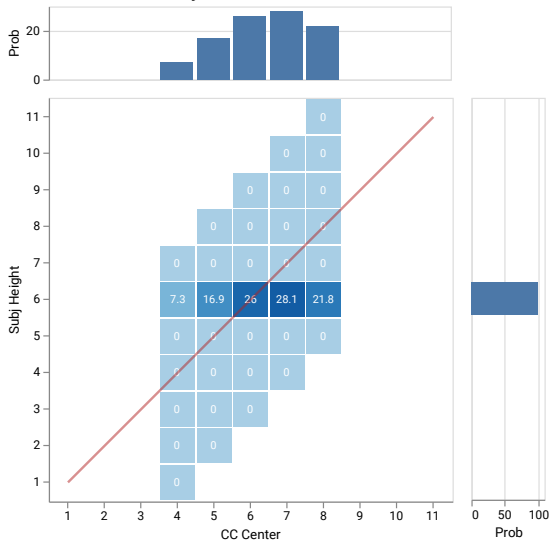
# degree demonstrative simulation: *Jane is that short*

L0 literal listener



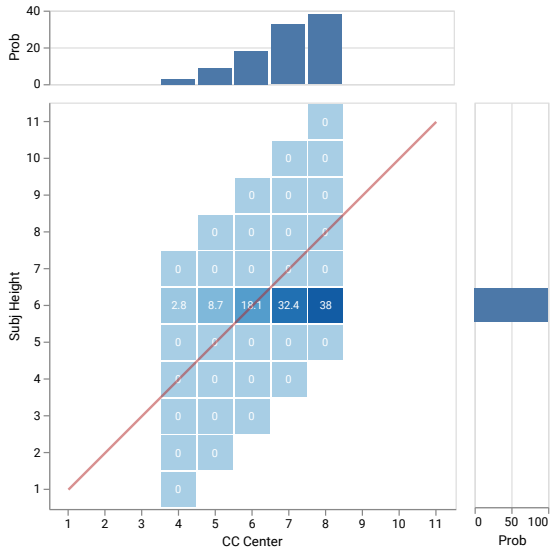
# degree demonstrative simulation: *Jane is that short*

L0 + lex. uncertainty



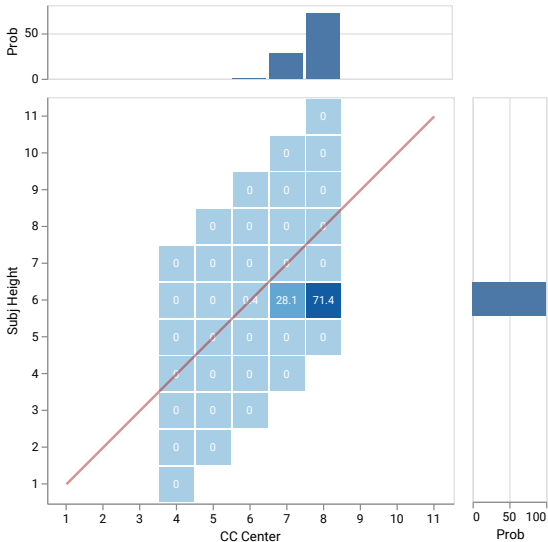
# degree demonstrative simulation: *Jane is that short*

## L1 pragmatic listener



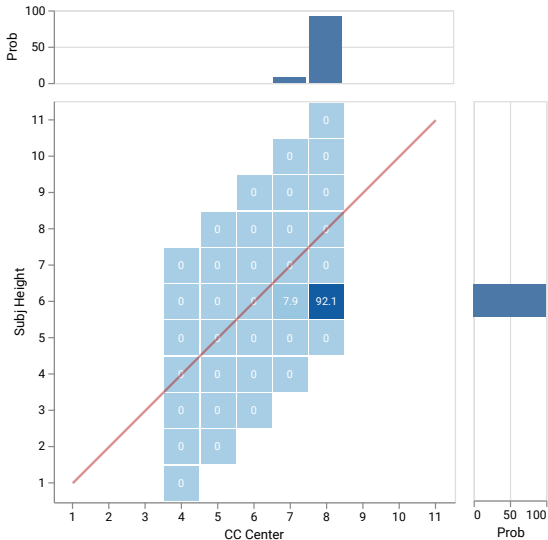
# degree demonstrative simulation: *Jane is that short*

L2 pragmatic listener



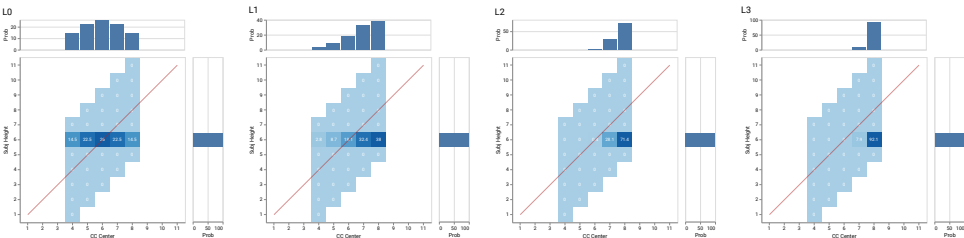
# degree demonstrative simulation: *Jane is that short*

## L3 pragmatic listener



## degree demonstrative summary: *Jane is that short*

- Again, since the lexica all entail that a "6-short person" has height 6, the distribution over Jane's height is certain



- But now, the various strengthenings lead ultimately to a pragmatic conclusion that 6 is almost certainly below the center
- evaluativity: **very strong**

## equative: semantic assumptions

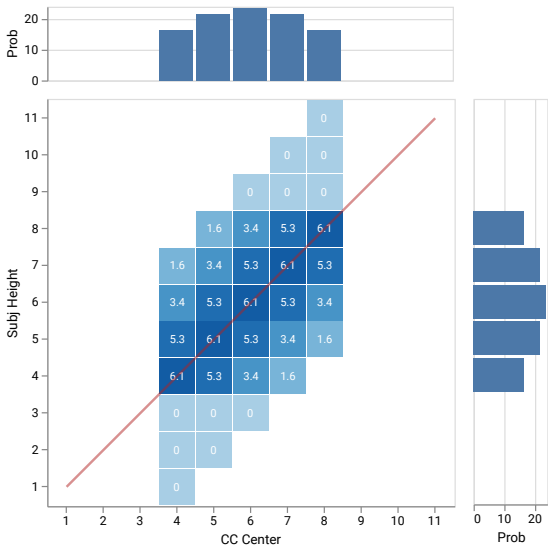
- *as tall/short as Keisha* both taken to entail that subject's height is equal to Keisha's height, which is unknown
- strengthened interpretations again place subject's height within various upper/lower percentiles of comp. class

	[[Jane is as tall as Keisha]]	[[Jane is as short as Keisha]]
⋮	⋮	⋮
$\mathcal{L}_{-1}$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \geq \mu_w - 1$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \leq \mu_w - 1$
$\mathcal{L}_0$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \geq \mu_w + 0$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \leq \mu_w + 0$
$\mathcal{L}_1$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \geq \mu_w + 1$	$\lambda w. \mathbf{ht}_w(j) = k \wedge k \leq \mu_w + 1$
⋮	⋮	⋮



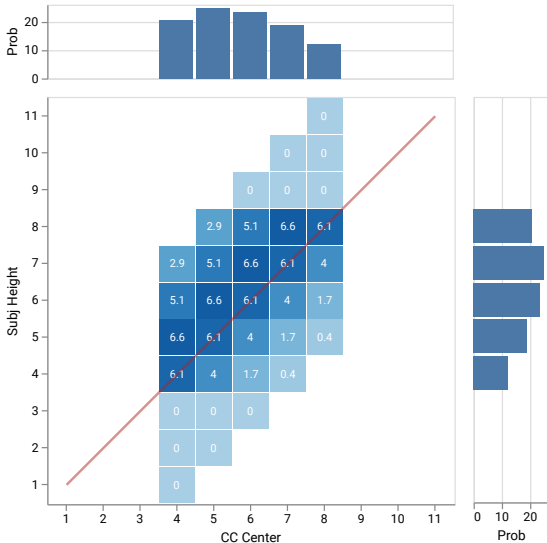
# equative simulation: *Jane is as tall as Keisha*

L0 literal listener



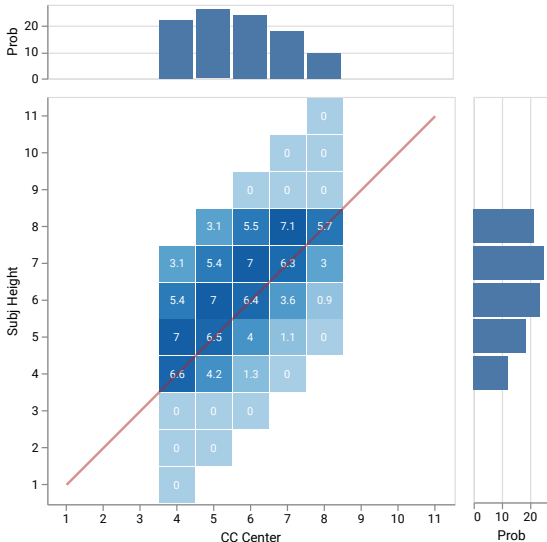
# equative simulation: *Jane is as tall as Keisha*

L0 + lex. uncertainty



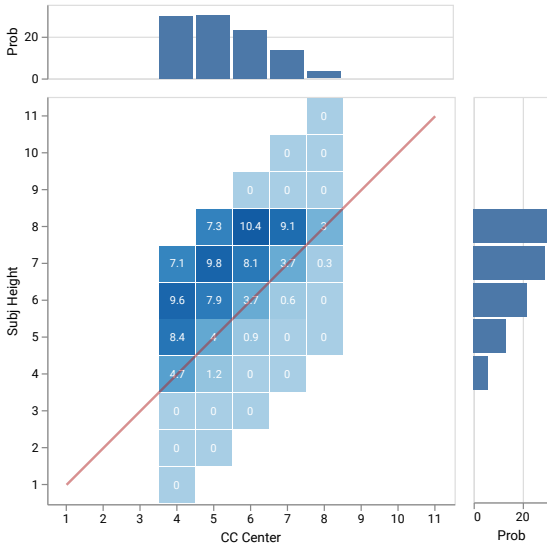
# equative simulation: *Jane is as tall as Keisha*

## L1 pragmatic listener



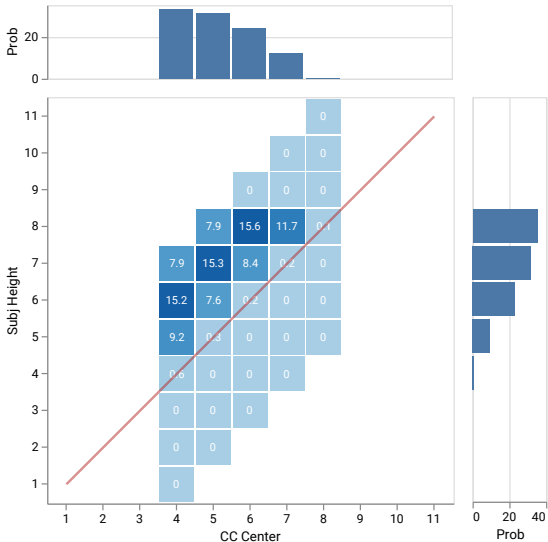
# equative simulation: *Jane is as tall as Keisha*

## L2 pragmatic listener



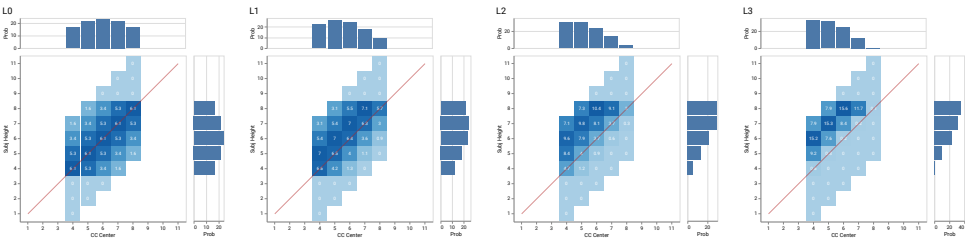
# equative simulation: *Jane is as tall as Keisha*

## L3 pragmatic listener



## equative summary: *Jane is as tall as Keisha*

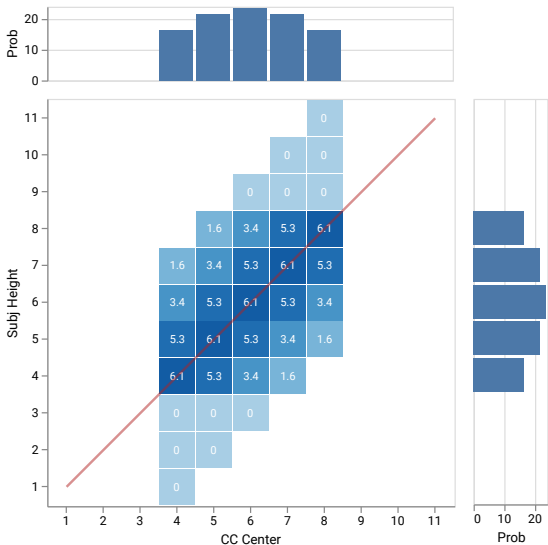
- Because Keisha's height is unknown, the literal (and pragmatic) posterior over worlds is spread out



- The various conceivable interpretations lead ultimately to a conclusion that Jane is probably a bit above the median
- evaluativity: **relatively weak**

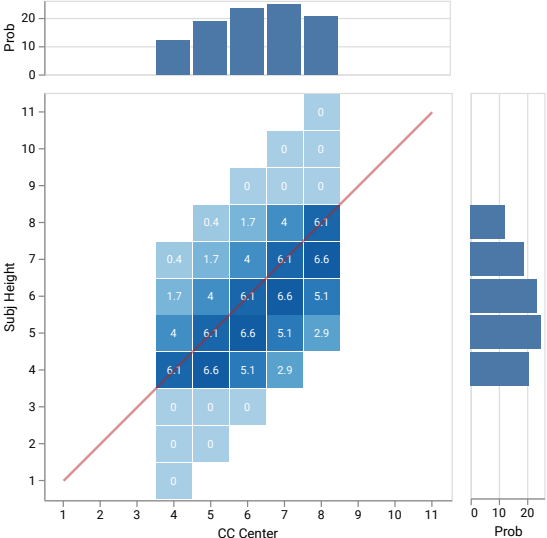
# equative simulation: *Jane is as short as Keisha*

L0 literal listener



# equative simulation: *Jane is as short as Keisha*

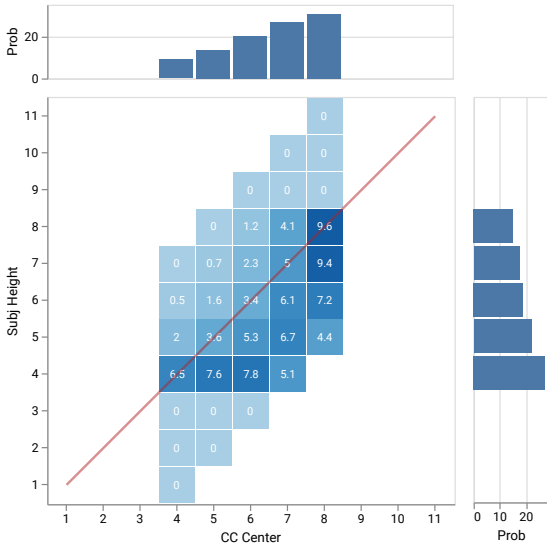
L0 + lex. uncertainty





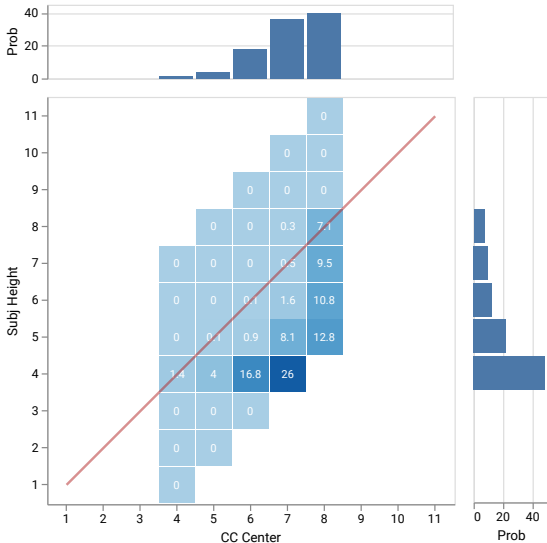
# equative simulation: *Jane is as short as Keisha*

L1 pragmatic listener



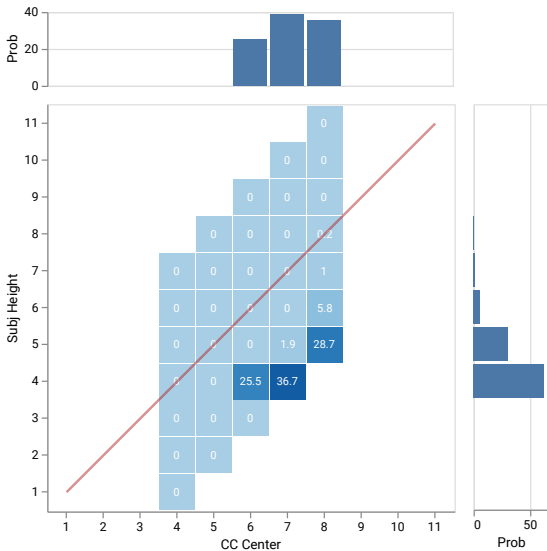
# equative simulation: *Jane is as short as Keisha*

## L2 pragmatic listener



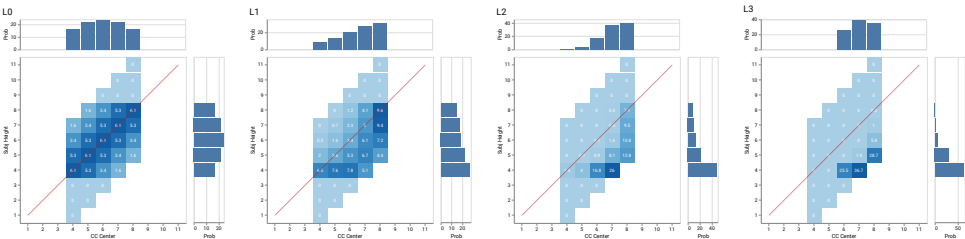
# equative simulation: *Jane is as short as Keisha*

## L3 pragmatic listener



## equative summary: *Jane is as short as Keisha*

- Again, since Keisha's height is unknown, the posterior probability remains distributed across many heights

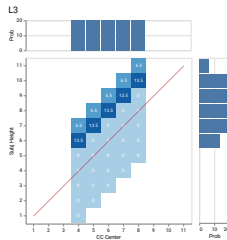
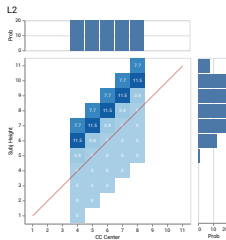
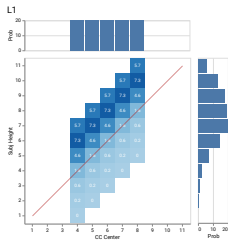
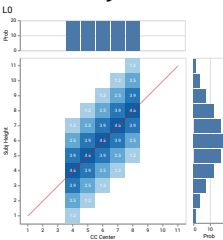


- But now, the various possible strengthenings lead to a pragmatic conclusion that Jane is almost certainly below the median
- evaluativity: **very strong**

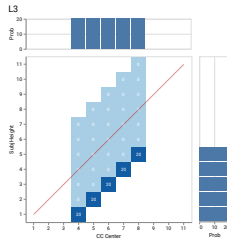
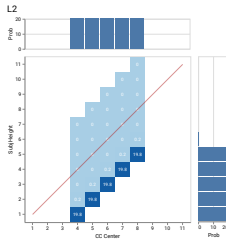
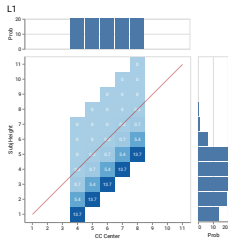
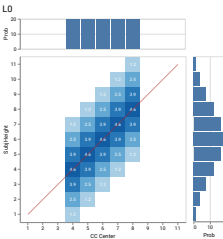
# results summary 1

positive form: **both antonyms evaluative**

- *Jane is tall*



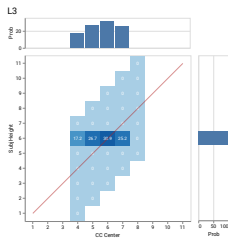
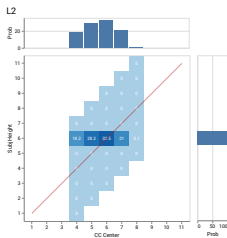
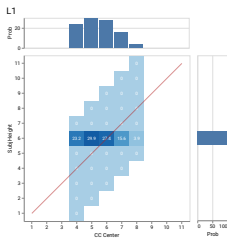
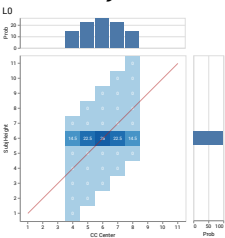
- *Jane is short*



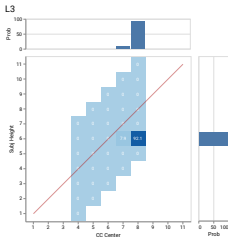
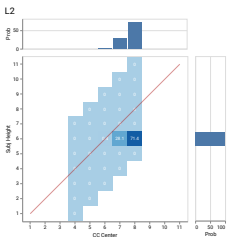
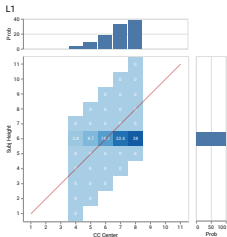
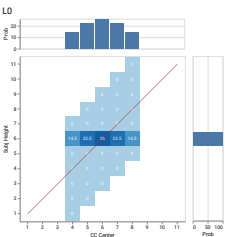
# results summary 2

degree demonstrative: **antonym-sensitive evaluativity**

- *Jane is that tall*



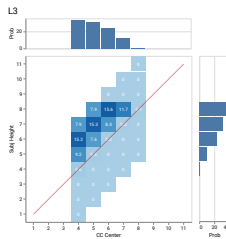
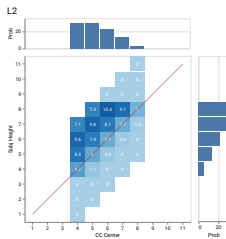
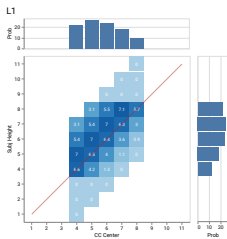
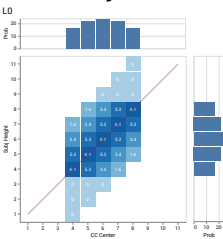
- *Jane is that short*



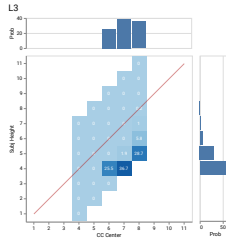
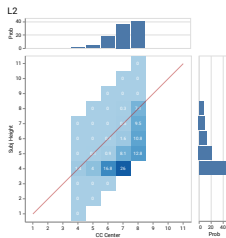
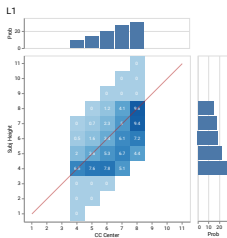
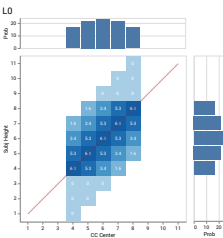
# results summary 3

## equative: **antonym-sensitive** evaluativity

- *Jane is as tall as Keisha*



- *Jane is as short as Keisha*



## conclusions

- evaluative inferences are not limited to positive constructions, so cannot be driven only by the need to fill in a degree argument
- in demonstrative and equative constructions, positive and negative antonyms compete semantically, so manner drives reasoning
- We adapted the schematic Bergen et al. (2016) model of lexical competition under semantic uncertainty to derive evaluativity inferences in particular, given suitable semantic entries
- canonical degree constructions are not just vague, they're context-sensitive: unsaturated degree parameters model the former (L&G 2013), and background distribution parameters model the latter (Barker 2002)
- in principle extendable to non-adjectival evaluative constructions (Rett 2015) as well as context-sensitive phenomena writ large



## selected references

- Barker, C. 2002. The dynamics of vagueness. *Linguistics and Philosophy* 25(1):1–36.
- Bergen, L., Levy, R., and Goodman, N. 2016. Pragmatic reasoning through semantic inference. *Semantics and Pragmatics* 9.
- Bierwisch, M. 1989. The semantics of gradation. In *Dimensional Adjectives: Grammatical Structure and Conceptual Interpretation*.
- Frank, M. and Goodman, N. 2012. Predicting pragmatic reasoning in language games. *Science* 336:998.
- Lassiter, D. and Goodman, N. 2013. Context, scale structure, and statistics in the interpretation of positive-form adjectives? In *SALT 23*.
- Lehrer, A. 1985. Markedness and antonymy. *Journal of Linguistics* 21(2): 397-429.
- Potts, C., Lassiter, D., Levy, R., and Frank, M. 2016. Embedded implicatures as pragmatic inferences under Compositional Lexical Uncertainty. *Journal of Semantics* 33:755–802.
- Rett, J. 2015. *The semantics of evaluativity*. Oxford University Press.
- Tessler, M. H., Lopez-Brau, M., and Goodman, N. D. 2017. Warm (for winter): Comparison class understanding in vague language. In *15th International Conference on Cognitive Modeling*.